ECM:

$$\nabla\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} -\hat{\alpha}\rho_1 & \rho_2 \\ -\hat{\alpha}\rho_2 & \rho_2 \end{bmatrix}\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \sum_i^n \begin{bmatrix} \beta_{x,i} & \beta_{y,i} \\ \gamma_{x,i} & \gamma_{y,i} \end{bmatrix} \nabla\begin{bmatrix} X_{t-i} \\ Y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

$$\begin{bmatrix} -\hat{\alpha} \\ -\hat{\alpha} \end{bmatrix}\begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix}$$

80

~~*~~ State-space model (SSM)

State-space models are useful tools for expressing dynamic systems that involve w/ unobserved state variables.

Two eqn.s:

(1) Measurement eqn (observation eqn.)
→ describes the relation btwn observed variables (data) and unobserved state variable.

(2) Transition eqn. (state eqn.)
→ describes the dynamics of the state variables. The transition equation has the form of a first-order difference equation in the state vector.

(1) $y_t = F' \underset{\sim}{x}_t + \varepsilon_t$

(2) $\underset{\sim}{x}_t = G \underset{\sim}{x}_{t-1} + \underset{\sim}{w}_t$

y_t = univariate time series; $t = 1, \cdots, T$.

$\underset{\sim}{x}_{t\,p\times 1}$ = state vector, $\underset{\sim}{x}_t \sim VAR(1)$.

$F_{p\times 1}$, $G_{p\times p}$ = coefficient matrices

$\varepsilon_t \sim N(0, \sigma^2)$

$\underset{\sim}{w}_t \sim N_p(\vec{0}, W)$.

e.g. $y_t \sim AR(2)$.

$\Phi_2(B) y_t = a_t$, $a_t \overset{iid}{\sim} N(0, \sigma^2)$.

(1) $y_t = F' \underset{\sim}{x}_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma^2)$ #1

(2) $\underset{\sim}{x}_t = G \underset{\sim}{x}_{t-1} + \underset{\sim}{w}_t \qquad \underset{\sim}{w}_t \sim N_p(\vec{0}, W)$.

$\underset{\sim}{x}_t \sim$ VAR(1) $\qquad W = E(\underset{\sim}{w}_t \underset{\sim}{w}_t')$

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t. \Rightarrow y_{t-1} = \phi_1 y_{t-2} + \phi_2 y_{t-3} + a_{t-1}$$

$$= \begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + a_t$$

$$\begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ y_{t-3} \end{bmatrix} + \begin{bmatrix} a_{t-1} \\ 0 \end{bmatrix}$$

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t.$$

WAY 1) Define $\underset{\sim}{x}_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$.

Then $y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$,

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} a_t \\ 0 \end{bmatrix}$$

WAY 2) Define $\underset{\sim}{x}_t = \begin{bmatrix} y_t \\ \phi_2 y_{t-1} \end{bmatrix}$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \phi_2 y_{t-1} \end{bmatrix},$$

$$\begin{bmatrix} y_t \\ \phi_2 y_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \phi_2 y_{t-2} \end{bmatrix} + \begin{bmatrix} a_t \\ 0 \end{bmatrix}.$$

e.g. $y_t \sim$ ARMA(1,1).

$$y_t - \phi_1 y_{t-1} = a_t + \theta_1 a_{t-1}$$

$$\Rightarrow y_t = \phi_1 y_{t-1} + a_t + \theta_1 a_{t-1}$$

Define $\underset{\sim}{x}_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$;

Then $y_t = \begin{bmatrix} 0 & \phi_1 \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} + a_t + \theta_1 a_{t-1}$,

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & 0 \\ & \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \\ \end{bmatrix}$$

$y_t \sim ARMA(1,1)$

$$y_t = \phi_1 y_{t-1} + \theta_1 a_{t-1} + a_t$$

$$\underset{\sim}{x_t} = \begin{bmatrix} y_t \\ \theta_1 a_t \end{bmatrix}, \quad \underset{\sim}{w_t} = \begin{bmatrix} a_t \\ \theta_1 a_t \end{bmatrix}$$

$$y_t = [1 \ \ 0] \begin{bmatrix} y_t \\ \theta_1 a_t \end{bmatrix}$$

$$\begin{bmatrix} y_t \\ \theta_1 a_t \end{bmatrix} = \begin{bmatrix} \phi_1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \theta_1 a_{t-1} \end{bmatrix} + \begin{bmatrix} a_t \\ \theta_1 a_t \end{bmatrix}$$

- Combining two SSM

$$y_t = F_1' \underset{\sim}{x_t} + F_2' \underset{\sim}{z_t} + \varepsilon_t,$$

$$\underset{\sim}{x_t} = G_1 \underset{\sim}{x_{t-1}} + \underset{\sim}{w_t}$$

$$\underset{\sim}{z_t} = G_2 \underset{\sim}{z_{t-1}} + \underset{\sim}{u_t}$$

Combined SSM: $y_t = [F_1' \ F_2'] \begin{bmatrix} \underset{\sim}{x_t} \\ \underset{\sim}{z_t} \end{bmatrix} + \varepsilon_t,$

$$\begin{bmatrix} \underset{\sim}{x_t} \\ \underset{\sim}{z_t} \end{bmatrix} = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} \begin{bmatrix} \underset{\sim}{x_{t-1}} \\ \underset{\sim}{z_{t-1}} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{w_t} \\ \underset{\sim}{u_t} \end{bmatrix}$$

- SSM for reg. model with AR(2) error.

$$y_t = \alpha + \beta f_t + \eta_t, \quad \eta_t \sim AR(2).$$

$$(\eta_t = \phi_1 \eta_{t-1} + \phi_2 \eta_{t-2} + a_t)$$

Let $\underset{\sim}{x_t} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$. Then $\underset{\sim}{x_t} = \underset{\sim}{x_{t-1}}$.

So $y_t = [1 \ \ f_t] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \eta_t$.

Also, $\underset{\sim}{z_t} = \begin{bmatrix} \eta_t \\ \phi_2 \eta_{t-1} \end{bmatrix} \Rightarrow \eta_t = [1 \ \ 0] \begin{bmatrix} \eta_t \\ \phi_2 \eta_{t-1} \end{bmatrix}$

and $\underset{\sim}{z_t} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \underset{\sim}{z_{t-1}} + \begin{bmatrix} a_t \\ 0 \end{bmatrix}$

$$y_t = \underbrace{[1 \ \ f_t]}_{F_1'} \underbrace{\begin{bmatrix}\alpha \\ \beta\end{bmatrix}}_{x_t} + \underbrace{[1 \ \ 0]}_{F_2'} \underbrace{\begin{bmatrix}\eta_t \\ \phi_2\eta_{t-1}\end{bmatrix}}_{z_t}$$

$$x_t = \begin{bmatrix}\alpha \\ \beta\end{bmatrix} = \underbrace{\begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}}_{G_1} \begin{bmatrix}\alpha \\ \beta\end{bmatrix} + \begin{bmatrix}0 \\ 0\end{bmatrix},$$

$$z_t = \begin{bmatrix}\eta_t \\ \phi_2\eta_{t-1}\end{bmatrix} = \underbrace{\begin{bmatrix}\phi_1 & 1 \\ \phi_2 & 0\end{bmatrix}}_{G_2} \begin{bmatrix}\eta_{t-1} \\ \phi_2\eta_{t-2}\end{bmatrix} + \begin{bmatrix}a_t \\ 0\end{bmatrix}$$

$$\Rightarrow y_t = [F_1' \ \ F_2'] \begin{bmatrix}x_t \\ z_t\end{bmatrix} = [1 \ \ f_t \ \ 1 \ \ 0] \begin{bmatrix}\alpha \\ \beta \\ \eta_t \\ \phi_2\eta_{t-1}\end{bmatrix}$$

$$\begin{bmatrix}x_t \\ z_t\end{bmatrix} = \begin{bmatrix}G_1 & 0 \\ 0 & G_2\end{bmatrix} \begin{bmatrix}x_{t-1} \\ z_{t-1}\end{bmatrix} + \begin{bmatrix}\vec{0} \\ a_t \\ 0\end{bmatrix}$$

$$\begin{bmatrix}\alpha \\ \beta \\ \eta_t \\ \phi_2\eta_{t-1}\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 0\end{bmatrix} \begin{bmatrix}\alpha \\ \beta \\ \eta_{t-1} \\ \phi_2\eta_{t-2}\end{bmatrix} + \begin{bmatrix}0 \\ 0 \\ a_t \\ 0\end{bmatrix}$$

STA355: Estimate the sampling dist. of $\hat{\theta}$ by sampling w/ replacement from the data and then computing estimates from these samples.

* Bootstrap

→ Replicating an experiment, by resampling from observed data

→ Used to estimate bias, se, percentiles, or other properties of a statistic

. General approach

→ Draw repeated samples w/ replacement

→ Since TS obs aren't iid, the method needs to be modified.

e.g. library(boot), bigcity

→ population (in 1000's) of 49 U.S. cities in 1920 (u) and 1930 (x).

→ 49 cities are among the biggest 196 cities in the U.S. in 1920. Randomly chosen.

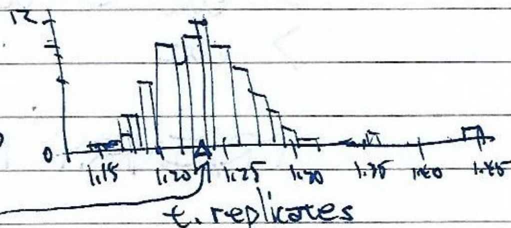
~~X~~ = U = pop. in ~~200~~ 1920
~~(U)~~ X = pop. in 1930

→ parameter of interest: $\theta = E(X)/E(U)$.
Let $\hat{\theta} = \bar{X}/\bar{U}$.
So how "uncertain" or "certain" are we about $\hat{\theta}$?

```
> dat <- bigcity[1:4,]
> dat
    u    x
1  138  143
2   93  104
3   61   69
4  179  260
> set.seed(1234)
> i <- sample(1:4, size=4, replace=T)
> i
[1] 1 3 3 3
> dat[i,]
      u    x
1    138  143
3     61   69
3.1   61   69
3.2   61   69
> ratio <- function(X) { mean(X[,2])/mean(X[,1])}
> n <- nrow(bigcity)
> n
[1] 49
> t0 <- ratio(bigcity)
> t0
[1] 1.239019
```

p5

```
> set.seed(1234)
> t.replicates <- replicate(1000, expr={
+ i <- sample(1:n, size=n, replace=T)
+ boots.obs <- bigcity[i,]
+ ratio(boots.obs)
+ })
> library(MASS)
> truehist(t.replicates)
> points(t0, 0, pch=17, col=2)
```



"▲" has appeared.

$$\hat{Var}(\hat{\theta}) = \frac{1}{R-1}\sum_j^R (\hat{\theta}^{(j)} - \bar{\hat{\theta}}^*)^2$$

$\bar{\hat{\theta}}^*$ is the arithmetic mean of the replicates.

```
> mean(t.replicates)
[1] 1.241886
> se.hat <- sd(t.replicates)
> se.hat
[1] 0.03635716
```

$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$.

And $\hat{Bias}(\hat{\theta}) = \bar{\hat{\theta}}^* - \hat{\theta}$.

```
> bias.hat <- mean(t.replicates) - t0
> bias.hat
[1] 0.002867242
```

- boot function (in boot package)

boot(data, statistic, R)

data: matrix or data.frame

statistic: name of a function to compute the statistic of interest

R: # of replicates to generate

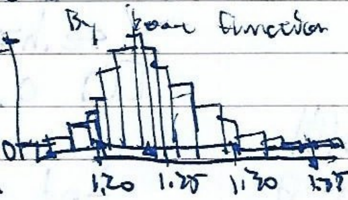
```
> ratio.boot <- function(X, i) {
+ y <- X[i,]
+ mean(y[,2])/mean(y[,1])
+ }
> out <- boot(bigcity, ratio.boot, 999)
> out

ORDINARY NONPARAMETRIC BOOTSTRAP

Call:
boot(data = bigcity, statistic = ratio.boot, R = 999)

Bootstrap Statistics:
      original    bias         std. error
t1*   1.239019    0.001209301  0.03499868

> truehist(out$t, main = "By boot function")
> points(t0, 0, pch=17, col=2)
> out$t0
[1] 1.239019   # explicit computation
> mean(out$t)
[1] 1.243081
> quantile( out$t, c(0.025, 0.975))
    2.5%       97.5%
  1.179312   1.317506
> quantile( out$t, c(0.025, 0.975), type=1) #edf
    2.5%       97.5%
  1.179143   1.317539
```



$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$

$\hat{F}(x) = \frac{\sum_{i=1}^{n} I(X_i \le x)}{n}$

"N" (norm)
Percentile CI (perc)
"t" (stud)
Basic Bootstrap (basic)
BCa CI (bca) ← "better" bootstrap CI

> boot.ci(out, type=c("norm", "perc", "basic", "bca"))

Level	Normal	Basic
95%	(1.166, 1.304)	(1.160, 1.299)

Level	Percentile	BCa
95%	(1.179, 1.318)	(1.179, 1.318)

— Bootstrapping regression

$y_t = \underset{\sim}{X_t}\underset{\sim}{\beta} + u_t$, $E(u_t|\underset{\sim}{X_t}) = 0$, $u_t \sim iid(0, \sigma^2)$.

n obs, k regressors.

Assumptions we make: ① independent errors?
② identically distributed errors?

· Types: Residual bootstrap
Parametric bootstrap
Wild bootstrap
Pair bootstrap

· What about bootstrap for dependent data?
→ Types: Parametric bootstrap
Sieve bootstrap ← popular approaches
Block bootstrap ← popular approaches
Carlstein (nonoverlapping)
~~Künsch~~ Künsch (overlapping)

· Comments on block bootstrap methods

$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$

— Bootstrap regression

n obs.
k regressors

$y_t = \underset{\sim}{X}_t \beta + u_t, \quad E(u_t \mid X_t) = 0, \quad u_t \sim iid(0, \sigma^2).$

Is u_t i? Is u_t id?

Types:
residual
parametric
wild
pair
(these all assume errors are indep.)

• Residual bootstrap

→ requires errors to be indep. of contemporaneous regressors and iid, but w/ minimal distributional assumptions

→ Steps:

① Use OLS to obtain $\hat{\beta}$ and $\hat{u}_t$.

② (Optional) Rescale residuals so that they have correct variance.

e.g. $\ddot{u}_t = \left(\frac{n}{n-k}\right)^{\frac{1}{2}} \hat{u}_t$.

③ Generate a typical observation of the bootstrap sample as

$y_t^* = \underset{\sim}{X}_t \hat{\beta} + u_t^*, \quad u_t^* \sim edf(\ddot{u}_t).$

u_t^*'s are often said to be resampled from $\ddot{u}_t$

• Parametric bootstrap

→ assume u_t follow a specific distribution, e.g. N.

→ Steps:

① Use OLS to obtain $\hat{\beta}$ and $\hat{u}_t$.

② Generate a typical observation using

$y_t^* = \underset{\sim}{X}_t \hat{\beta} + u_t^*, \quad u_t^* \overset{iid}{\sim} N(0, s^2).$

s^2 = sample var of $\hat{u}_t$.

$S_{yy} = Y'(I - \frac{1}{n}J)Y$

$SSReg = Y'(H - \frac{1}{n}J)Y$

$RSS = Y'(I-H)Y$
$= \hat{e}'\hat{e}$

$\hat{e} = Y - \hat{Y} = (I-H)Y$

$\cancel{\hat{\beta} = X'(X'X)^{-1}X}$

$\hat{\beta} = (X'X)^{-1}X'Y$

$H = X(X'X)^{-1}X'$

e.g. 

$s^2 = \frac{RSS}{n-(p+1)}$

$\widehat{Var}(\hat{\beta}) = s^2(X'X)^{-1}$

• Wild bootstrap

→ is designed to handle heteroskedasticity in reg. models.
  (non-constant variance among residuals)

→ data generating process (DGP):

$y_t^* = \underset{\sim}{X}_t \hat{\beta} + f(\hat{u}_t) v_t^*$ $\quad v_t^* \sim dist(0,1)$,

e.g. $v_t^* \sim Bern(0.5)$ (-1 or 1).

$f(\hat{u}_t)$ is a transformation of the t^{th} residual $\hat{u}_t$.

e.g. $f(\hat{u}_t) = \frac{\hat{u}_t}{\sqrt{1-h_{ii}}}$.

• Pair bootstrap

→ resample from the matrix with typical row $[y_t \; \underset{\sim}{X}_t]$.
We no longer condition on the X_t since each bootstrap sample now has a different $\underset{\sim}{X}$ matrix.
A typical observation of the bootstrap sample is $[y_t^* \; \underset{\sim}{X}_t^*]$.

→ Comments:

① the pairs of bootstrap is valid even when the errors display heteroskedasticity of unknown form

② it works even for dynamic models.

③ pairs bootstrap can be applied to an enormous range of models.

④ in the case of multivariate models, we can combine the pairs and residual bootstraps: organize residuals as a matrix and apply the pairs bootstrap to its rows. This preserves cross-equation correlations.

→ Deficiencies:

(1) The bootstrap DGP does not impose H_0's restrictions on β.

(2) Compared to residual bootstrap (under validity) and wild bootstrap, pairs bootstrap does not yield very accurate results.

$\cancel{x_t = \phi_1 x_{t-1} + a_t + \alpha}$

$\cancel{x_t^2 = \phi_1 x_t x_{t-1} + a_t x_t + \alpha x_t}$

$\cancel{\gamma_0 = \phi_1 \gamma_1 + \sigma_a^2 + \alpha^2 \gamma_0}$

$\cancel{\gamma_1 = \phi_1 \gamma_0 + \alpha \gamma_0}$

— Bootstrap for dependent data

Types: Parametric, Sieve, Block

→ Resampling breaks up any dependence and is therefore inappropriate for dependent data.

→ Two popular approaches: sieve and block

- Parametric bootstrapping

Consider stationary $x_t \sim AR(1)$.

$\mu = \phi_1 \mu + \alpha$

$(1-\phi_1)\mu = \alpha$

$\mu = \frac{\alpha}{1-\phi_1}$

$$x_t = \phi_1 x_{t-1} + a_t + \alpha, \quad a_t \overset{iid}{\sim} N(0, \sigma_a^2) \quad \text{— (1)}$$

$$\text{(Unconditional)} \quad x_t \sim N\left(\frac{\alpha}{1-\phi_1}, \frac{\sigma_a^2}{1-\phi_1^2}\right) \quad \text{— (2)}$$

Why

$$x_t = \alpha + \phi_1 x_{t-1} + a_t$$

$$x_t x_{t-k} = \alpha x_{t-k} + \phi_1 x_{t-1} x_{t-k} + a_t x_{t-k}$$

$$\gamma_k = \alpha \underbrace{E(x_{t-k})}_{=0} + \phi_1 \gamma_{k-1} + E(a_t x_{t-k})$$

$\cancel{k=0}$ $\Rightarrow \gamma_k = \phi_1 \gamma_{k-1} + E(a_t x_{t-k})$.

$k=0 \Rightarrow E(a_t x_t)$

$= E(a_t(\cancel{\alpha} + \phi_1 x_{t-1} + a_t))$

$= 0 + 0 + \sigma_a^2 = \sigma_a^2$.

$\Rightarrow \gamma_0 = \phi_1 \gamma_1 + \sigma_a^2$

$k=1 \Rightarrow \gamma_1 = \phi_1 \gamma_0$

$\Rightarrow \gamma_0 = \phi_1(\phi_1 \gamma_0) + \sigma_a^2$

$= \phi_1^2 \gamma_0 + \sigma_a^2$

$\Rightarrow (1-\phi_1^2)\gamma_0 = \sigma_a^2$

$\Rightarrow \underline{\gamma_0 = \frac{\sigma_a^2}{1-\phi_1^2}}$ //

$x_t = \phi_1 x_{t-1} + \alpha + a_t$

$\frac{\alpha}{1-\phi_1} = \phi_1 \frac{\alpha}{1-\phi_1} + \alpha$

$\frac{1}{1-\phi_1} = \frac{\phi_1}{1-\phi_1} + 1$

$$\text{(Conditional)} \quad X_t | X_{t-1} \sim N(\phi_1 X_{t-1}, \sigma_a^2) \quad \text{— (3)}$$

→ The (unconditional) simulation procedure may be summarized as:

① Simulate X_0 by drawing a random # from (2)

② Simulate $X_1 = \alpha + \phi_1 X_0 + a_t$.

③ Simulate $X_t = \alpha + \phi_1 X_{t-1} + a_t$ recursively.

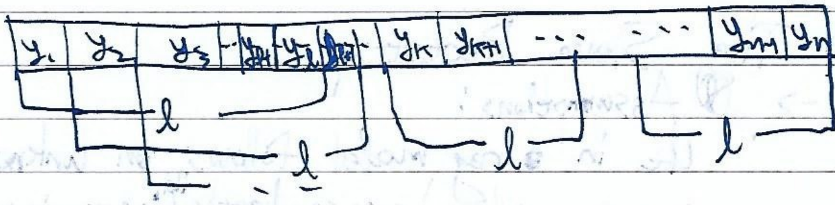
$$y_t = \underset{\sim}{X}_t \beta + u_t$$

- The Sieve Bootstrap.
  - → Assumptions:
    u_t in a reg model follows an unknown, stationary process w/ homoscedastic, iid innovations.
  - → Approximate such process by AR(p) where p is chosen by some sort of model selection criteria like AIC or BIC, or by sequential testing (p-1 vs. p)
  - → Steps:
    1. Estimate the model to obtain residuals $\hat{u}_t$.
    2. Estimate AR(p) model: $\hat{u}_t = \sum_{i=1}^{p} \phi_i \hat{u}_{t-i} + \varepsilon_t$. (1)
    3. Generate bootstrap error terms
       $u_t^* = \sum_{i=1}^{p} \hat{\phi}_i u_{t-i}^* + \varepsilon_t^*$, (2)
       where ε_t^*'s are resampled from the (rescaled) residuals from (1).
    4. Generate the bootstrap data according to
       $y_t^* = \underset{\sim}{X}_t \hat{\beta} + u_t^*$, where u_t^* is from (2).

- MBB (the moving block bootstrap)
  - → Carlstein : bootstrapping blocks of observations, rather than individual observations; nonoverlapping blocks
  - → Künsch : MBB, applicable to stationary time series data; blocks of observations are overlapping.

  ~~Steps~~ Idea

  ① Divide the data of n observations into blocks of length l and select b of these blocks ~~with replacement~~ by resampling with replacement all the possible blocks.



Carlstein	Künsch
Block 1: $(y_1, y_2, \cdots, y_l)$	Block 1: $(y_1, \cdots, y_l)$
Block 2: $(y_{l+1}, \cdots, y_{2l})$	Block 2: $(y_2, \cdots, y_{l+1})$
⋮	⋮
Block k: $(y_{(k-1)l+1}, \cdots, y_{kl})$	Block k: $(y_k, \cdots, y_{l+k-1})$
⋮	⋮
Block last (b): $(y_{(b-1)l+1}, \cdots, y_n)$	Block last (b): $(y_b, \cdots, y_n)$
$\Rightarrow$ # of blocks: $b = \frac{n}{l}$. So $n = lb$.	$\Rightarrow$ # of blocks: $b = n-l+1$. So $n = l+b-1$.

Block 3 $(y_{2l+1}, \cdots, y_{3l})$

(l = size of a block, b = # of blocks)

e.g. $x_t = \{\boxed{3}\boxed{6}\boxed{7}\boxed{2}\boxed{1}\boxed{5}\}$.
$l=3$, $n=6$.

Carlstein $\Rightarrow b=2$	Künsch: $b=4$
$\boxed{3\,6\,7} = ☆_1$, $\boxed{2\,1\,5} = ☆_2$	$\boxed{3\,6\,7} = □_1$, $\boxed{6\,7\,2} = □_2$, $\boxed{7\,2\,1} = □_3$, $\boxed{2\,1\,5} = □_4$
Draw a sample of two blocks w/ replacement: possible outcomes:	Draw a sample of two blocks w/ replacement: possible outcomes:
$\Omega = \{(☆_1, ☆_1), (☆_1, ☆_2), (☆_2, ☆_1), (☆_2, ☆_2)\}$	$\Omega = \{(□_1, □_1), \cdots, (□_4, □_4)$. $\lvert\Omega\rvert = 16$.

~~P(¬☆2 in the second | ☆1 was 1st)~~

$P_c(\neg ☆_2$ in 2^{nd} | $☆_1$ was $1^{st})$
$= 1/2 = \underline{0.5}$.

$P_k(\neg □_4$ in 2^{nd} | $□_1$ was $1^{st})$
$= 1/4 = \underline{0.25}$.

Problems w/ MBB

① Even if $\{X_t\}$ is stationary, the pseudo time series generated from $\{X_t\}$ by MBB is not stationary.

Possible solution: let $\ell \sim Geom(p) \ni p$.

② The mean of MBB, $\bar{x}_n^*$, is biased:
$E(\bar{x}_n^* \mid x_1, \cdots, x_n) - \bar{x}_n \neq 0$.

③ MBB estimator of the variance of $\sqrt{n} \cdot \bar{x}_n$ is also biased; instead of $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$,
we should use $\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{\ell}\left[(x_i - \bar{x})^2 + \sum_{k=1}^{\ell-1}\sum_{i=1}^{n-k}(x_i - \bar{x}_n)(x_{i+k} - \bar{x}_n)\right]$.

Optimal ℓ

① Carlstein < Künsch

② $\ell \sim Geom(p)$, or $\ell = \frac{1}{p}$ ($E(\ell)$).
→ less sensitive to the choice of p than the application of MBB is to the choice of ℓ.

Comments

- Block bootstrap methods divide the quantities that are being resampled, which might be either rescaled residuals or $[y, X]$ pairs, into blocks of ℓ consecutive observations. We then resample the blocks.
- Overlapping blocks are better
- ℓ fixed is better than ℓ variable.
- Choice of ℓ is critical; it must ↑ as n↑. Often, $\ell \propto n^{1/3}$.

· l too small ⇒ cannot mimic original.
Dependence is broken.
· l too long ⇒ bootstrap samples are not random enough.
· Block-of-blocks bootstrap is the analog of the pairs bootstrap for dynamic models.

Steps

Consider $y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} + \gamma y_{t-1} + u_t$, $u_t \overset{iid}{\sim} dist(0, \sigma^2)$.

Define $\underset{\sim}{Z}_t = [y_t, y_{t-1}, \underset{\sim}{X}_t]$.

Construct length l overlapping blocks:

Block 1 : $(\underset{\sim}{Z}_1, \cdots, \underset{\sim}{Z}_l)$
Block 2 : $(\underset{\sim}{Z}_2, \cdots, \underset{\sim}{Z}_{l+1})$
⋮
Block k : $(\underset{\sim}{Z}_k, \cdots, \underset{\sim}{Z}_{l+k-1})$
⋮
Block b (last) : $(\underset{\sim}{Z}_b, \cdots, \underset{\sim}{Z}_n)$.

$(n = l+b-1 \Rightarrow b = n-l+1$ blocks in total).

→ BofB bootstrap works w/ non-constant variance feature as well as serial correlation.
→ Generally, Block bootstraps offer higher-order accuracy than asymptotic methods, but only by a modest extent.
→ Block bootstraps can yield more reliable standard errors.